Reference: *Biol. Bull.* **197:** 198–206. (October 1999)

Origin of Insulin Receptor-Like Tyrosine Kinases in Marine Sponges

ALEXANDER SKOROKHOD[1,2], VERA GAMULIN[3], DIETMAR GUNDACKER[1], VADIM KAVSAN[2], ISABEL M. MÜLLER[1], AND WERNER E. G. MÜLLER[1,*]

[1] *Institut für Physiologische Chemie, Abteilung Angewandte Molekularbiologie, Universität, Duesbergweg 6, D-55099 Mainz, Germany;* [2] *Ukrainian Academy of Sciences, Department of Biosynthesis of Nucleic Acids, Institute of Molecular Biology and Genetics, 252627 Kiev, Ukraine; and* [3] *Institute Rudjer Boskovic, Department of Molecular Genetics, 10000 Zagreb, Croatia*

Abstract. One autapomorphic character restricted to all Metazoa including Porifera [sponges] is the existence of transmembrane receptor tyrosine kinases (RTKs). In this study we screened for molecules from one subfamily within the superfamily of the insulin receptors. The subfamily includes the insulin receptors (InsR), the insulin-like growth factor I receptors, and the InsR-related receptors—all found in vertebrates—as well as the InsR-homolog from *Drosophila melanogaster*. cDNAs encoding putative InsRs were isolated from the hexactinellid sponge *Aphrocallistes vastus*, the demosponge *Suberites domuncula*, and the calcareous sponge *Sycon raphanus*. Phylogenetic analyses of the catalytic domains of the putative RTKs showed that the sponge polypeptides must be grouped with the InsRs. The relationships revealed that all sponge sequences fall into one branch of this group, whereas related sequences from mammals (human, mouse, and rat), insects and molluscs, and polypeptides from one cephalochordate, fall together into a second branch. We have concluded that (*i*) the InsR-like molecules evolved in sponges prior to the "Cambrian Explosion" and contributed to the rapid appearance of the higher metazoan phyla; (*ii*) the sponges constitute a monophyletic taxon, and (*iii*) epidermal growth factor (EGF)-like domains are present in sponges, which allows the insertion of this domain into potential receptor and matrix molecules.

Received 13 January 1999; accepted 26 July 1999.

To whom correspondence should be addressed. E-mail: WMUELLER@mail.UNI-Mainz.DE

The sequences reported here are deposited in the EMBL/GenBank data base: *Suberites domuncula*, cDNA for InsR-like molecule, SDINR: accession number Y17880; *Aphrocallistes vastus*, cDNA for InsR-like molecule AVtNR: accession number Y17881; *Sycon raphanus*, cDNA for type 1 InsR-like molecule, SRtNR1: accession number Y17877; *Sycon raphanus*, cDNA for type 2 InsR-like molecule, SRtNR2: accession number Y17878; *Sycon raphanus*, cDNA for type 3 InsR-like molecule, SRtNR3: accession number Y17879.

Abbreviations: aa, amino acid; kb, kilobase; nt, nucleotide; InsR, insulin receptor; ORF, open reading frame; EGF, epidermal growth factor; IGF-I-R, insulin growth factor I receptor; PTK, protein tyrosine kinase; RTK, receptor tyrosine kinase; TK, tyrosine kinase

Introduction

The Porifera [sponges] are the oldest metazoan phylum; they existed 40 to 50 million years prior to the onset of the "Cambrian Explosion" (Valentine *et al.*, 1996), the time of main divergence of metazoan phyla (Valentine, 1994). Highly conserved amino acid (aa) sequences in sponges indicate that the Porifera share one common ancestor with other metazoan phyla (Müller *et al.*, 1994; also see Müller, 1995, 1997, and 1998). These sequences include those (*i*) for transmembrane receptors, *e.g.*, transmembrane tyrosine kinase [TK] receptors [RTKs] (Müller and Schäcke, 1996); (*ii*) for transmembrane adhesion molecules, *e.g.*, the integrins (Pancer *et al.*, 1997a); and (*iii*) for G-protein linked transmembrane receptors for signaling molecules, *e.g.*, the metabotropic glutamate receptor (Perovic *et al.*, 1999). Additional sequences from homeodomain transcription factors show that the transcriptional control of gene expression in the oldest Metazoa is similar to that of the most recent phyla (Seimiya *et al.*, 1994; Richelle-Maurer *et al.*, 1998; Coutinho *et al.*, 1998). One metazoan autapomorphic character restricted to Porifera is the presence of high telomerase activity in all (or almost all) cells, including somatic cells (Koziol *et al.*, 1998).

The discovery that sponges contain transmembrane (Schäcke *et al.*, 1994a), cytoplasmic (Ottilie *et al.*, 1992), and nuclear TKs (Cetkovic *et al.*, 1998) suggests that the

signaling system in these animals is sophisticated enough to respond to peptide growth factors and to cell adhesion (Müller and Müller, 1999). The catalytic domain of the RTKs is related to that of the cytoplasmic protein tyrosine kinases [PTKs] and the Ser/Thr kinases (Hanks and Hunter, 1995; Kruse *et al.*, 1997). The catalytic domain of the TKs is subdivided into 12 smaller subdomains, the first eight of which are most highly conserved (Hardie and Hanks, 1995). In addition to the characteristic tyrosine protein kinase-specific active-site signature, the previously described catalytic domain of the RTK from the demosponge *Geodia cydonium* contains no further site that marks this molecule as belonging to a specific class of RTKs (Schäcke *et al.*, 1994a). In this study, we have demonstrated for the first time that one distinct subfamily of the RTKs is already present in all three classes of Porifera, and that it contains the TK class II signature with the consensus pattern D-[LIV]-Y-x_3-Y-Y-R (PC/GENE, 1995 [Prosite]). By choosing appropriate primers for the polymerase chain reaction, sequences were obtained from sponges that must be grouped with the insulin receptors [InsRs] of vertebrates (Ullrich *et al.*, 1985), the insulin-like growth factor I receptors [IGF-I-Rs] of vertebrates (Ullrich *et al.*, 1986), the InsR-related receptors of vertebrates (Shier and Watt, 1989), and the InsR homolog from *Drosophila melanogaster* (Fernandez *et al.*, 1995). These molecules are all members of the class II RTKs which display, within subdomain VII, the following consensus for InsRs, GF-I-Rs, and InsR-homologs R-D-[IV]-Y-E-[TS]-D-Y (Hardie and Hanks, 1995).

Here we present the TK domains of InsR-(like) molecules that have been isolated from the hexactinellid sponge *Aphrocallistes vastus*, the demosponge *Suberites domuncula*, and the calcareous sponge *Sycon raphanus*. From *S. raphanus*, three full-length clones from the InsR-like molecules are given. All of the sequences were used for phylogenetic analyses. These revealed that the sponge InsR-like molecules are statistically significantly distinct from the related molecules of higher Metazoa, and allowed an assessment of the evolutionary order in which the three classes of Porifera appeared.

Materials and Methods

Materials

Restriction endonucleases and other enzymes for recombinant DNA techniques and vectors were obtained from Stratagene (La Jolla, CA; USA), QIAGEN (Hilden; Germany), Boehringer Mannheim (Mannheim; Germany), GibcoBRL (Grand Island, NY; USA), Amersham (Buckinghamshire; UK), USB (Cleveland, OH; USA), DUPONT (Bad Homburg; Germany), Epicentre Technologies (Madison, WI; USA), and Promega (Madison, WI; USA). *Taq* DNA polymerase, DIG [digoxigenin] DNA labeling kit, DIG-11-dUTP, anti-DIG AP Fab fragments, and CDP [disodium 2-chloro-5-(4-methoxyspiro{1,2-dioxetane-3,2′-(5′-chloro)-tricyclo[3.3.1.1^{3,7}]decan}-4-yl)phenyl phosphate] were from Boehringer Mannheim (Mannheim; Germany).

Sponges

Live specimens of *Sycon raphanus* [Schmidt] (Porifera, Calcarea, Calcaronea, Leucosoleniida, Sycettidae) and *Suberites domuncula* [Olivi] (Porifera, Demospongiae, Tetractinomorpha, Hadromerida, Suberitidae) were collected from the Adriatic Sea near Rovinj (Croatia). The specimens of *Aphrocallistes vastus* [Schulze] (Porifera, Hexactinellida, Hexasterophora, Hexactinosida, Aphrocallistidae) were collected from Saanich Inlet and Barkley Sound, British Columbia (Canada) by scuba diving. They were a gift of Dr. Sally P. Leys (Department of Biology, University of Victoria, P.O. Box 1700, Victoria, BC, Canada). The material was immediately frozen in liquid nitrogen until use.

Construction of cDNA library from A. vastus

Total RNA was extracted from sponge tissue, and polyadenylated mRNA was isolated from total RNA as already described (Pfeifer *et al.*, 1993a and b). cDNA was prepared with a ZAP Express cDNA synthesis kit. The cDNA library of *A. vastus* was prepared in Hybri ZAPII (Stratagene) and packaged *in vitro* with the MaxPlax Packaging Extract (Epicentre Technologies). The library contained approximately 2.4×10^6 independent plaque forming units (pfu); the amplified library was stored at 4°C.

Screening and isolation of the cDNAs encoding InsR-like molecules

The complete cDNAs as well as those encoding the catalytic domains were cloned by the polymerase chain reaction (PCR) from the *A. vastus* cDNA library (see above), the *S. domuncula* (Kruse *et al.*, 1997), or the *S. raphanus* cDNA libraries (Kruse *et al.*, 1997). The degenerate sense primer 5′-TTYGGIATGGTITAYGARGG-3′ (Y = pyrimidine, R = purine, I = inosine) and the downstream primer (anti sense) 5′-TARTARTCIGTYTCRTADATRTC-3′ were designed against the conserved regions of TK subdomain I (FGMVYEG) and TK subdomain VII (DIYETDY) of InsRs as well as IGF-I-Rs from mammalian species; these regions are different from the corresponding protein kinases of other classes (Scavo *et al.*, 1991). The two primers define a 470–490 bp long sequence encoding part of the TK catalytic domain (Fig. 1). The PCR was carried out using a GeneAmp 9600 thermal cycler (Perkin Elmer), with an initial denaturation at 95°C for 3 min, then 35 amplification cycles each at 95°C for 30 s, 50°C for 45 s, 72°C for 1.5 min, and a final extension step at 74°C for 10 min. The reaction mixture of 50 μl included 20 pmol of the respective

degenerate primer and 10 pmol of the primer T7 (Stratagene), 200 μM of each nucleotide, 1 μl of the respective cDNA libraries, buffer, and 2.5 units of *Taq* DNA polymerase. The expected amplified products were purified and concentrated using Geneclean Spin Kit and directly ligated into pGEM-T vector. After isolation and purification, the plasmid DNAs were sequenced with an automatic DNA sequenator [Li-Cor 4200].

The TK catalytic domains of the three *S. raphanus* InsR-like molecules were used and completed by both 5′- and 3′-RACE, using the kits "5′-" and "3′-RACE System" to full-length cDNAs.

Sequence analyses

Sequences were analyzed using PC/GENE, release 14.0, from IntelliGenetics, Mountain View, CA (USA). Similarity searches and sequence retrieval were performed *via* the e-mail servers at the European Bioinformatics Institute, Hinxton Hall, UK (BLITZ and FASTA), and the National Institutes of Health, Bethesda, MD, USA (BLAST). The phylogenetic tree was constructed from an aa alignment by the neighbor-joining method (Saitou and Nei, 1987) applying the PHYLIP package version 3.5c program (Felsenstein, 1993). The degree of support for internal branches was further assessed by bootstrapping. The distance matrix was calculated as described (Dayhoff *et al.*, 1978). Multiple alignments were performed with CLUSTAL W version 1.6 (Thompson *et al.*, 1994) and their graphic presentations by the program GeneDoc (Nicholas and Nicholas, 1997).

Northern blot

RNA was extracted from liquid-nitrogen-pulverized sponge tissue with TRIzol Reagent (GibcoBRL) as recommended by the manufacturer. Total RNA (1 μg) was electrophoresed through formaldehyde/agarose gel and blotted onto Hybond N^+ membrane following the manufacturer's instructions (Amersham). Hybridization experiments were performed with the probes *SRINR1*, *SRINR2*, or *SRINR3* [≈600 bp segments] from *S. raphanus*. These probes were labeled with DIG-11-dUTP by the DIG DNA labeling kit. Hybridization was performed with the anti sense DIG-labeled probes at 42°C overnight using 50% formamide containing 5×SSC, 2% blocking reagent [Boehringer], 7% [w/v] SDS, and 0.1% [w/v] N-lauroylsarcosine, following the instructions of the manufacturer [Boehringer]. After washing, DIG-labeled nucleic acid was detected with anti-DIG Fab fragments [conjugated to alkaline phosphatase] and visualized by a chemiluminescence technique using CDP, the chemiluminescence substrate for alkaline phosphatase, according to the instructions of the manufacturer [Boehringer].

Results

Cloning and sequencing the cDNAs encoding the InsR-like molecules

The *S. domuncula* nt sequence, *SDINR*, is 491 nt long and has a potential open reading frame [ORF] of 489 bases encoding a deduced protein sequence of 163 aa residues. The sequence from *A. vastus*, *AVINR*, is 490 nt long with an ORF of 489 nt (163 aa).

Three putative sequences of InsR-like molecules were isolated from the cDNA library of *S. raphanus*. The cDNA for type 1 InsR, *SRINR1*, is 2026 nt long with an ORF of 1848 nt encoding a putative sequence of 616 aa (Figs. 1 and 2); type 2 InsR, *SRINR2*, is 2150 nt long with an ORF of 1842 nt (614 aa); and type 3 InsR (Fig. 2), *SRINR3*, is 1433 nt long with an ORF of 1368 nt (456 aa) (Fig. 2). Northern blot analyses were performed with these *S. raphanus* cDNA probes. One band each of approximately 2.2 kb (type 1), 2.3 kb (type 2), and 1.6 kb (type 3) were obtained, confirming that the full-length cDNAs were isolated (Fig. 3).

Deduced aa sequences of the catalytic domains of the putative sponge InsRs

The deduced aa sequences of the catalytic domains of the InsR-like sequences between subdomains I to VII have been aligned (Fig. 1). The borders of subdomains I to VII (according to Hardie and Hanks, 1995), could be defined for all sponge sequences unequivocally (Fig. 1). Specific sites and sequence characteristics were also present as outlined earlier (Müller and Schäcke, 1996): in subdomain I, the ATP-binding site [consensus: GxGxxGxV; but in the hexactinellid INR_AV sequence G is replaced by R]; within subdomain II, the residue Lys in the consensus VAxK, which is required for kinase activity; within subdomain VIb: the aa D [Asp] and N [Asn] as well as in subdomain VII: the DFG tripeptide is present. The DFG segment has been implicated in ATP binding (Hanks *et al.*, 1988) and represents the most conserved portion within the catalytic domain. The tyrosine residue (Y) in subdomain VII (aa no. 180 of the catalytic domain, with respect to the *G. cydonium* RTK) undergoes phosphorylation and is the tyrosine kinase phosphorylation site. Signatures within subdomains VIII, IX, X, and XI are generally less well conserved. Therefore, the PCR-based sequencing was restricted to the part within subdomains I to VII. The TK-specific active-site signature, D-L-A-T/A-R-N, characteristic for both vertebrate and invertebrate TKs (Ottilie *et al.*, 1992; Hanks *et al.*, 1988) is found in subdomain VIb. Within the subdomain VII the signature for the TK class II receptors with the consensus pattern is found, D-[LIV]-Y-x_3-Y-Y-R (PC/GENE, 1995 [Prosite]).

The PCR primers were chosen to identify, in sponges, those catalytic domains of class II RTKs that share the

highest similarity to InsRs, IGF-I-Rs, to InsR-related receptors, and to the insulin receptor homolog from *D. melanogaster* (Fernandez *et al.*, 1995). These receptors have the consensus within the class II signature of R-D-[IV]-Y-E-[TS]-D-Y (Hardie and Hanks, 1995). As seen in Figure 1, this consensus is, as expected, present in all sponge sequences; therefore the sequences from the demosponge *S. domuncula*, the calcareous sponge *S. raphanus*, and the hexactinellid sponge *A. vastus* were termed InsR-like molecules.

Complete aa sequences of the InsR-like sequences from S. raphanus

Three cDNAs encoding complete putative InsR-like sequences from *S. raphanus* have been isolated from the library. The sequences are termed type 1, *SRINR1*, type 2, *SRINR2*, and type 3, *SRINR3*, InsR-like molecules. The putative 616 aa sequence INR_SR1 (deduced from SRINR1) has a calculated M_r of 69,477; INR_SR2 of 614 aa has an M_r of 69,213, and INR_SR3 of 456 aa has an M_r of 51,259.

The sequence INR_SR2 was selected for the analysis given here. The transmembrane segment, determined according to the program "RAOARGOS" (PC/GENE, 1995) ranges from aa_{181} to aa_{196}. The intracellular domain is, as in other RTKs (Hardie and Hanks, 1995), divided into a juxtamembrane domain (aa_{197} to aa_{242}) and the catalytic domain [TK domain] (aa_{243} to aa_{521}) (Fig. 2). The catalytic domain is subdivided into 12 subdomains and contains the characteristic TK-specific active-site signature and the RTK class II signature (see above); in addition, the putative ATP-binding site (Hanks *et al.*, 1988) is present (Fig. 2).

The extracellular domain contains one calcium-binding, epidermal growth factor receptor [EGF]-like domain that reads D-x-N-E-C^1-D-x_5-C^2-D-E-C^3-Q-N-C^4-x-N-x_6-C^5-x-N-x_3-C^6-D; it is located from aa_{129} to aa_{162} [the Cys residues are numbered consecutively]. This EGF-like domain consists of six Cys residues, flanked by aa with carbonyl oxygen atoms, which are arranged slightly differently from those found in molecules from higher Metazoa. In particular, the Cys^4 and Cys^5 are separated by more than one aa (Bork *et al.*, 1996). Furthermore, an incomplete EGF-like domain is present from aa_{49} to aa_{128}. The two other types of InsR-like molecules from *S. raphanus* also have two EGF-like domains, and they are similarly arranged. This finding is the first demonstration that EGF-like domains are present in the lowest metazoan phylum. Until now, this domain, which is widely found in vertebrate receptors—*e.g.*, mammalian epidermal growth factor receptors (Geer *et al.*, 1994) and matrix proteins like fibulin (Pan *et al.*, 1993)—has only been identified among invertebrates in *Caenorhabditis elegans* (Campbell and Bork, 1993).

Phylogenetic analyses

When the deduced aa sequences of the TK catalytic domains from the three sponge species were analyzed using the programs BLITZ, FASTA, and BLAST, they displayed highest similarity to the polypeptides from both invertebrates and vertebrates. Among invertebrates, these domains were most similar to the insulin-like receptors from the insects *Aedes aegypti* and *D. melanogaster*, as well as to the insulin receptor of the mollusc *Aplysia californica*. In addition an InsR-homolog sequence isolated, so far, from one cephalochordate, *Branchiostoma lanceolatum*, as well as InsRs, IGF-I-Rs, or InsR-homologs from selected vertebrates (human, mouse and rat) were highly similar to the sponge sequences. They share about 40%–45% of identical aa and about 60%–65% of similar aa (including identical aa) with the selected corresponding molecules. Taking only the sponge sequences, the sequence from *S. raphanus*, type 1, is identical in 67% of the aa (similarity of 78%) within the catalytic domain with *A. vastus* and in 69% (79%) with *S. domuncula*. The finding that the three sequences obtained from *S. raphanus* differ considerably from each other is interesting; type 1 shares only 75% identical aa (similarity 86%) with type 2 and only 79% identical aa (similarity 88%) with type 3.

The phylogenetic tree was constructed and rooted with the sequence of the catalytic domain of the Fes/FER non-receptor TK domain from *S. raphanus* (Cetkovic *et al.*, 1998; Fig. 4A). All sequences used were cut for the alignment to obtain the 12 subdomains, comprising approximately 300 aa. All of the sponge sequences fall into one branch of the tree, whereas the selected sequences of InsRs, IGF-I-Rs, or InsR-related sequences from invertebrates and vertebrates are grouped together into a second one. This relationship is statistically very robust as analyzed by bootstrapping. Hence, support for monophyly of Porifera can be deduced. In consequence, the presented findings, based on the data obtained with the catalytic domains of the InsR-like molecules from sponges, shed new light on the assumed uncertain position of sponges as reviewed by Rodrigo *et al.* (1994). In addition, the data given do not support earlier notions which suggested that the phylum Porifera might be paraphyletic (Cavalier-Smith *et al.*, 1996).

Rate of evolution of the catalytic domains of sponge InsR-like molecules

Use of our data collected on the percentage of aa identity among the polypeptide sequences from the different sponge species on one side and the sponge sequences in comparison to those from higher metazoan allows a relative approach to determining the time of divergence of the sponge classes from a common ancestor. This estimation, which is based on the number of point mutations per 100 aa within given polypeptides, might reflect the time of divergence of two

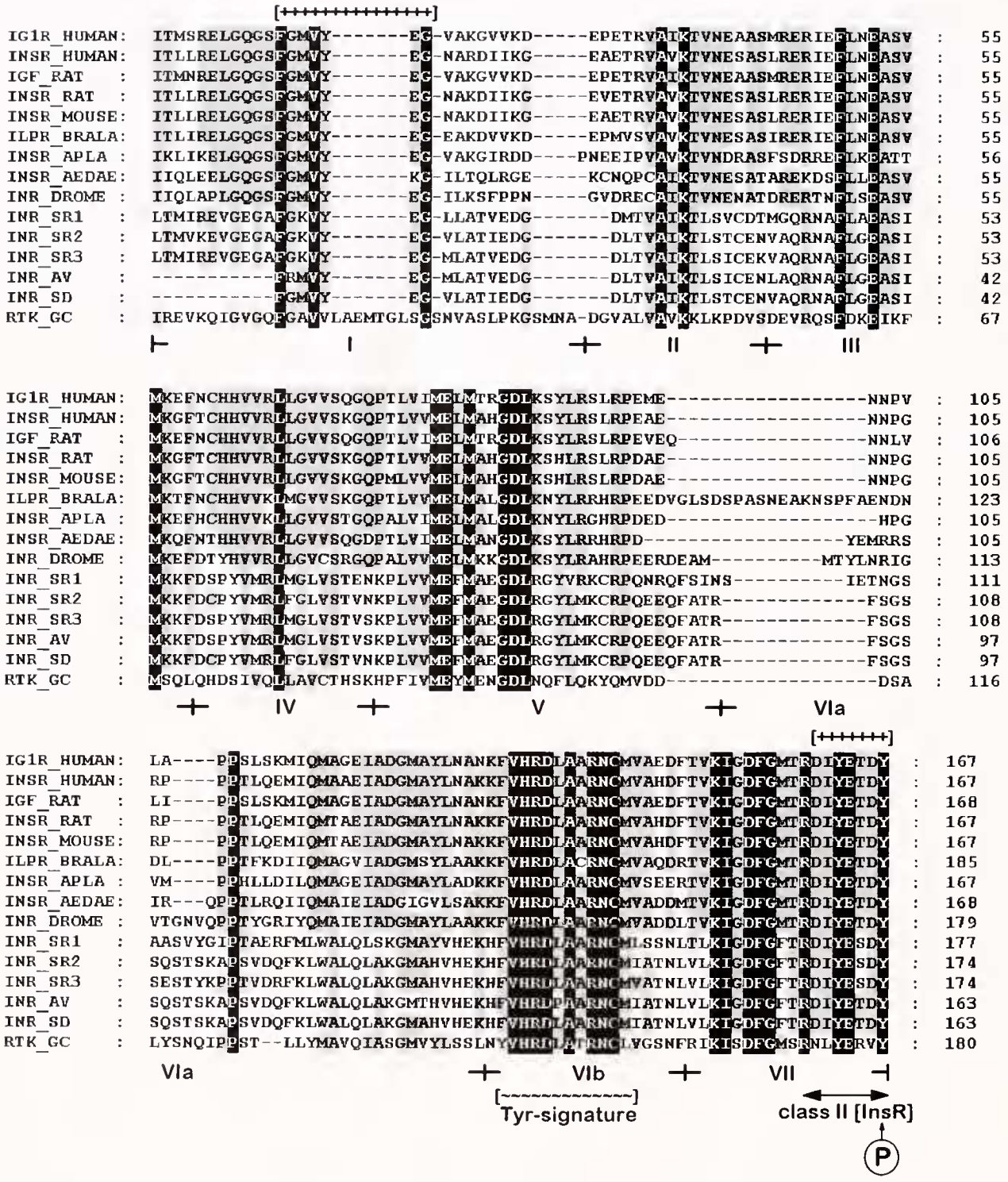

Figure 1. Alignment of the catalytic domains of the insulin-like receptors as well as of the related sequences within the class II tyrosine kinase receptors (RTKs). The deduced amino acid sequences of InsR homologs from the polypeptides of the three classes of the phylum Porifera (sponges) were aligned with the related sequences for three invertebrates (mosquito, fruit fly, and mollusc); one cephalochordate (amphioxus); and the three vertebrates (human, mouse, rat). In addition, the RTK domain from the sponge *Geodia cydonium* was used for the comparison. Specific sequence codes are identified below. The borders of the subdomains [I to VII] are shown in the figure below the sequences; the nomenclature follows that of Hardie and Hanks (1995). Also marked are the tyrosine protein kinase specific active-site signature (Tyr-signature [~~~~]), the TK class II

taxa. The evolutionary rates—expressed as k_{aa}-values—vary between different proteins (Zuckerkandl and Pauling, 1965; Kimura, 1983; Li *et al.*, 1987). In a previous study, the galectin protein from the sponge *G. cydonium* (Pfeifer *et al.*, 1993b) was calculated to have an estimated evolutionary rate of 0.97×10^{-9} aa substitutions/site/year (Hirabayashi and Kasai, 1993); a value of 1.24×10^{-9} was calculated for the RTK from *G. cydonium* (Schäcke *et al.*, 1994b) from the same animal.

Dating based on the molecular clock is inaccurate because its rate often varies. If we accept this insecurity, reject the estimated evolutionary rate from sponge genes, and accept the one calculated from the time of protostome-deuterostome divergence—700 MYA (Dayhoff 1978)—we can postulate the time of separation of the sponges from the common metazoan ancestor, as follows. If we take the calculated k_{aa}-value for the human to *D. melanogaste* (0.46) as a reference for the protostome-deuterostome split, then the hexactinellid sponge *A. vastus* branched off 1400 MYA (k_{aa}-value of 0.92), followed by the demosponge *S. domuncula* 1300 (k_{aa}-value of 0.84) and the calcareous sponge *S. raphanus* 1200 MYA for type 1 and 2 (k_{aa}-value of 0.80) and for type 3 1100 MYA (k_{aa}-value of 0.77). Recent fossil data show (Li *et al.*, 1998) that sponges existed in much their present form 580 MYA (Fig. 4B).

Discussion

We have shown that all three classes of the phylum Porifera express molecules related to InsR; and these molecules display, in their extracellular domains, EGF-like sequences (as shown here for *S. raphanus*). This finding implies that animals of the lowest metazoan phylum already contain growth factor receptors that allow them to react to nutrient cues and also to neighboring, individual cells, with a complex intracellular signaling reaction. The InsR-homologs, which are putative transmembrane receptors, presumably allow the transduction of signals through the cellular membrane. Usually signaling by RTKs involves ligand-mediated receptor dimerization (Geer *et al.*, 1994), a process that has not yet been studied in Porifera. InsRs, IGF-1-Rs, and InsR-related receptors or InsR-homologs of higher metazoan taxa do not contain, in their extracellular loops, EGF-like domains, but rather cysteine-rich regions (Geer *et al.*, 1994). This finding underlines again previous findings, that most polypeptides deduced from the cDNA sequences of sponges are assembled by an unusually large variety of modules. For one example, the putative sponge aggregation receptor is composed of scavenger receptor cysteine-rich domains as well as of short consensus repeats (Pancer *et al.*, 1997b; Blumbach *et al.*, 1998) in a structural complexity not known in higher Metazoa.

From the evolutionary point of view, the present contribution makes three points. First, it establishes that molecules similar to the InsR-homologs have evolved prior to the "Cambrian Explosion." Suga *et al.* (1997) suggested that most of the PTK subfamilies, including InsRs, diverged by domain shufflings, together with gene duplications before the diploblast-triplobast split. As a result of recent findings that the Porifera already existed before this event (Li *et al.*, 1998), we can assume that this class of key molecules, involved in the complex network of intracellular signaling, could have been one major driving force that allowed the

---

signature [specific for InsRs and related sequences] (←class II→ [InsR]) as well as the TK phosphorylation site (P). The positions of the primers are indicated ([+ + + +]) above the sequences. Identical aa residues in all 15 sequences are shown in white-on-black, and residues conserved in at least eight sequences are shaded.

Vertebrates

IGIR_HUMAN = Human insulin-like growth factor 1 receptor precursor (XO4434)
INSR_HUMAN = Human InsR precursor (PO6213)
INSR_MOUSE = InsR precursor house mouse (*Mus musculus;* P15208),
IGF_RAT = IGF-1-R1 receptor precursor rat (*Rattus norvegicus;* A33837)
INSR_RAT = InsR precursor from rat (*R. norvegicus;* P15127)

Cephalochordate

ILPR_BRALA = Insulin-like peptide receptor precursor amphioxus (*Branchiostoma lanceolatum;* O02466)

Invertebrates

INSR_AEDAE = Insulin-like receptor precursor mosquito (*Aedes aegypti;* Q93105)
INR_DROME = InsR homolog fruit fly (*Drosophila melanogaster;* U28136)
INSR_APLA = InsR the mollusc (*Aplysia californica;* 1587845)

Sponges

INR_SD = Insulin-like receptor demosponge *Suberites domuncula*
INR_SR1 = Insulin-like receptor calcareous sponge *Sycon raphanus* type 1
INR_SR2 = Insulin-like receptor calcareous sponge *S. raphanus* type 2
INR_SR3 = Insulin-like receptor calcareous sponge *S. raphanus* type 3
INR_AV = Insulin-like receptor hexactinellid sponge *Aphrocallistes vastus*
RTK_GC = Receptor tyrosine kinase demosponge *Geodia cydonium* (X77528)

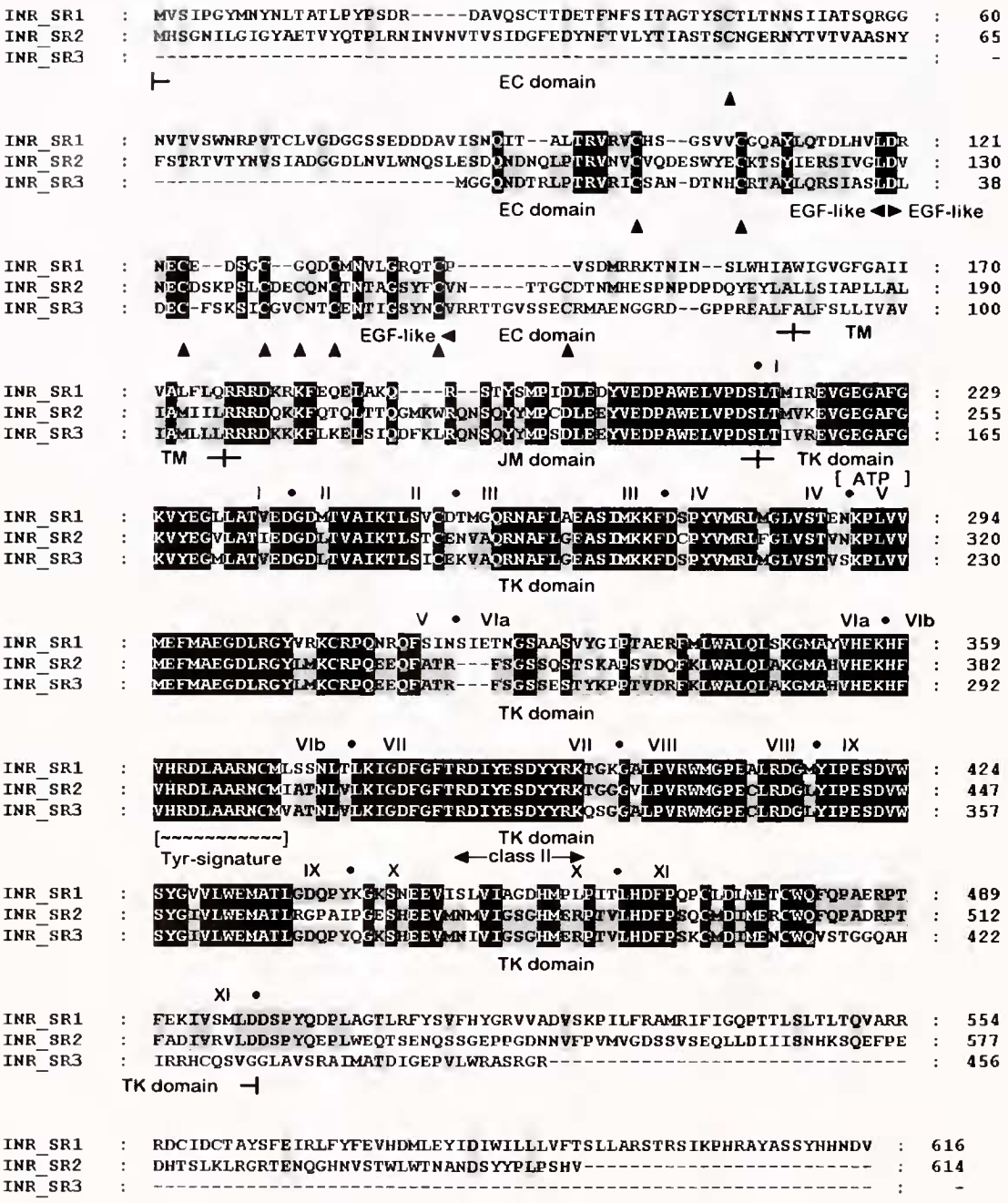

Figure 2. Alignment of the deduced aa sequences of the InsR-like sequences from *Sycon raphanus* type 1 (INR_SR1), type 2 (INR_SR2), and type 3 (INR_SR3). The four segments of the sequences are the extracellular domains (EC domain); the transmembrane segments (TM); and the two intracellular domains, the juxtamembrane domain (JM domain) and the catalytic domain (TK domain). The TK domain is subdivided into the 12 subdomains (above the sequence) with the characteristic TK-specific active-site signature (Tyr-signature [~~~~~]) and RTK class II signature (←class II→) (below the sequence); in addition, the putative ATP-binding-site (ATP) is marked. In the extracellular domain, the conserved Cys residues (arrowhead) of the two EGF-like domains (EGF-like) are indicated. Identical amino acid residues are shown in white-on-black type, and residues conserved in at least two sequences are shaded.

other metazoan phyla to arise. Second, the phylogenetic analyses confirm that, based on the autapomorphic character for Metazoa, the RTKs, sponges as a taxon are monophyletic; the Hexactinellida have been calculated to be the oldest class, followed by the Demospongia and finally by the Calcarea. Third, EGF-like domains are already present

Figure 3. Northern blot analyses to determine the sizes of the transcripts of the mRNA encoding the *Sycon raphanus* InsR-like molecules (*SRINR*) type 1 (*SRINR1*), type 2 (*SRINR2*), and type 3 (*SRINR3*). RNA was prepared from sponge tissue and 1 μg each was subjected to analysis. Molecular masses of marker RNAs, which were run in parallel, are given on the right (in kilobytes).

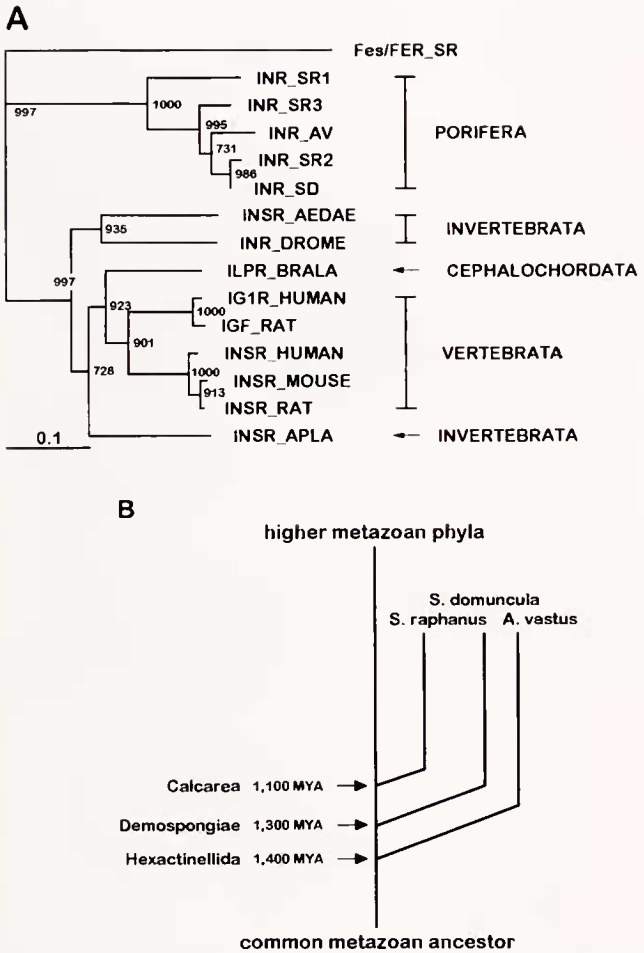

Figure 4. (A) Rooted phylogenetic tree of the catalytic domains of the sequences listed in Figure 1. The numbers at the nodes refer to the levels of confidence as determined by bootstrap analysis. The scale bar indicates an evolutionary distance of 0.1 amino acid substitution per position in the sequence. The catalytic domain of the Fes/FER nonreceptor TK domain from *Sycon raphanus* (Fes/FER_SR, Y17051) was used as the outgroup sequence. (B) Proposed branching order of the three classes of the phylum Porifera (Hexactinellida, Demospongiae, and Calcarea) from a common metazoan ancestor. The dates of the approximate times of divergence are indicated.

in sponges, where they were inserted into potential cell surface receptors and also into matrix molecules.

Acknowledgments

This work was supported by grants from the Deutsche Forschungsgemeinschaft [Mü 348/12-1] and from the International Human Frontier Science Program [W. E. G. Müller; RG-333/96-M].

Literature Cited

Blumbach, B., Z. Pancer, B. Diehl-Seifert, R. Steffen, J. Münkner, I. Müller, and W. E. G. Müller. 1998. The putative sponge aggregation receptor: isolation and characterization of a molecule composed of scavenger receptor cysteine-rich domains and short consensus repeats. *J. Cell. Sci.* **111:** 2635–2644.

Bork, P., A. K. Downing, B. Kieffer, and I. D. Campbell. 1996. Structure and distribution of modules in extracellular proteins. *Q. Rev. Biophys.* **29:** 119–167.

Campbell, I. D., and P. Bork. 1993. Epidermal growth factor-like molecules. *Curr. Opin. Struct. Biol.* **3:** 385–392.

Cavalier-Smith, T., M. T. E. P. Allsopp, E. E. Chao, N. Boury-Esnault, and J. Vacelet. 1996. Sponge phylogeny, animal monophyly, and the origin of the nervous system: 18S rRNA evidence. *Can. J. Zool.* **74:** 2031–2045.

Cetkovic, H., I. M. Müller, W. E. G. Müller, and V. Gamulin. 1998. Characterization and phylogenetic analysis of a cDNA encoding the Fes/FER related, non-receptor protein-tyrosine kinase in the marine sponge *Sycon raphanus*. *Gene* **216:** 77–84.

Coutinho, C. C., J. Seack, G. Van de Vyver, R. Borojevic, and W. E. G. Müller. 1998. Origin of metazoan bodyplan: characterization and functional testing of the homeobox gene *EmH-3* from the freshwater sponge *Ephydatia muelleri* in mouse 3T3 cells. *Biol. Chem. Hoppe-Seyler* **379:** 1243–1251.

Dayhoff, M. O. 1978. Survey of new data and computer methods of analysis. Pp. 1–8 in *Atlas of Protein Sequence and Structure*, Vol. 5, suppl. 3, M. Dayhoff, ed. National Biomedical Research Foundation, Washington, DC.

Dayhoff, M. O., R. M. Schwartz, and B. C. Orcutt. 1978. A model of evolutionary change in protein. Pp. 345–352 in *Atlas of Protein Sequence and Structure*, Vol. 5, suppl. 3, M. Dayhoff, ed. National Biomedical Research Foundation, Washington, DC.

Felsenstein, J. 1993. PHYLIP (Phylogeny Inference Package), ver. 3.5. University of Washington, Seattle.

Fernandez, R., D. Tabarini, N. Azpiazu, M. Frasch, and J. Schlessinger. 1995. The *Drosophila* insulin receptor homologue: a gene essential for embryonic development encodes two receptor isoforms with different signaling potential. *EMBO J.* **14:** 101–112.

Geer, P. v. d., T. Hunter, and R. A. Lindberg. 1994. Receptor protein-tyrosine kinases and their signal transduction pathways. *Annu. Rev. Cell Biol.* **10:** 251–337.

Hanks, S. K., and T. Hunter. 1995. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.* **9:** 579–596.

Hanks, S. K., M. A. Quin, and T. Hunter. 1988. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* **241:** 42–52.

Hardie, G., and S. Hanks. 1995. *The Protein Kinase FactsBook: Protein-tyrosine Kinases.* Academic Press, London.

Hirabayashi, J., and K. Kasai. 1993. The family of metazoan metal-independent β-galactoside-binding lectins: structure, function and molecular evolution. *Glycobiology* **3:** 297–304.

Kimura, M. 1983. *The Neutral Theory of Molecular Evolution.* Cambridge University Press, Cambridge.

Koziol, C., R. Borojevic, R. Steffen, and W. E. G. Müller. 1998. Sponges (Porifera) model systems to study the shift from immortal- to senescent-somatic cells: the telomerase activity in somatic cells. *Mech. Ageing Dev.* **100:** 107–120.

Kruse, M., I. M. Müller, and W. E. G. Müller. 1997. Early evolution of metazoan serine/threonine and tyrosine kinases: identification of selected kinases in marine sponges. *Mol. Biol. Evol.* **14:** 1326–1334.

Li, W. H., M. Tanimura, and P. M. Sharp. 1987. An evaluation of the molecular clock hypothesis using mammalian DNA sequences. *J. Mol. Evol.* **25:** 330–342.

Li, C. W., J. Y. Chen, and T. E. Hua. 1998. Precambrian sponges with cellular structures. *Science* **279:** 879–882.

Müller, W. E. G. 1995. Molecular phylogeny of Metazoa (animals): monophyletic origin. *Naturwissenschaften* **82:** 321–329.

Müller, W. E. G. 1997. Origin of metazoan adhesion molecules and adhesion receptors as deduced from their cDNA analyses from the marine sponge *Geodia cydonium*. *Cell Tissue Res.* **289:** 383–395.

Müller, W. E. G. 1998. Origin of Metazoa: sponges as living fossils. *Naturwissenschaften* **85:** 11–25.

Müller, W. E. G., and I. M. Müller. 1999. Primordial endocrinology in the lowest invertebrates: The Porifera. in *Reproductive Biology of Invertebrates,* A. Dorn, ed. J. Wiley & Sons, Chichester. In press.

Müller, W. E. G., and H. Schäcke. 1996. Characterization of the receptor protein-tyrosine kinase gene from the marine sponge *Geodia cydonium*. *Prog. Mol. Subcell. Biol.* **17:** 183–208.

Müller, W. E. G., I. M. Müller, and V. Gamulin. 1994. On the monophyletic evolution of the Metazoa. *Braz. J. Med. Biol. Res.* **27:** 2083–2096.

Nicholas, K. B., and H. B. Nicholas, Jr. 1997. *GeneDoc: a Tool for Editing and Annotating Multiple Sequence Alignments.* Distributed by the author. Version 1.1.004. www.cris.co./~ketchupTgenedoc.shtml.

Ottilie, S., F. Raulf, A. Barnekow, G. Hannig, and M. Schartl. 1992. Multiple *src*-related genes, *srk1-4,* in the fresh water sponge *Spongilla lacustris*. *Oncogene* **7:** 1625–1630.

Pan, T. C., T. Sasaki, R. Z. Zhang, R. Fassler, R. Timpl, and M. L. Chu. 1993. Structure and expression of fibulin-2, a novel extracellular matrix protein with multiple EGF-like repeats and consensus motifs for calcium binding. *J. Cell Biol.* **123:** 1269–1277.

Pancer, Z., M. Kruse, I. Müller, and W. E. G. Müller. 1997a. On the origin of adhesion receptors of metazoa: cloning of the integrin α subunit cDNA from the sponge *Geodia cydonium*. *Mol. Biol. Evol.* **14:** 391–398.

Pancer, Z., J. Münkner, I. Müller, and W. E. G. Müller. 1997b. A novel member of an ancient superfamily: sponge (*Geodia cydonium,* Porifera) putative protein that features scavenger receptor cysteine-rich repeats. *Gene* **193:** 211–218.

PC/GENE, Data Banks CD-ROM. 1995. [EMBL, Swiss Prot]; Release 14.0. IntelliGenetics, Inc. Mountain View, CA.

Perovic, S., I. Prokic, A. Krasko, I. M. Müller, and W. E. G. Müller. 1999. Origin of neuronal receptors in Metazoa: cloning of a metabotropic glutamate-like receptor from the marine sponge *Geodia cydonium*. *Cell Tissue Res.* **296:** 395–404.

Pfeifer, K., M. Haasemann, V. Gamulin, H. Bretting, F. Fahrenholz, and W. E. G. Müller. 1993a. S-type lectins occur also in invertebrates: high conservation of the carbohydrate recognition domain in the lectin genes from the marine sponge *Geodia cydonium*. *Glycobiology* **3:** 179–184.

Pfeifer, K., W. Frank, H. C. Schröder, V. Gamulin, B. Rinkevich, R. Batel, I. M. Müller, and W. E. G. Müller. 1993b. Cloning of the polyubiquitin cDNA from the marine sponge *Geodia cydonium* and its preferential expression during reaggregation of cells. *J. Cell Sci.* **106:** 545–554.

Richelle-Maurer, E., G. Van de Vyver, S. Visser, and C. C. Coutinho. 1998. Homeobox-containing genes in freshwater sponges: characterization, expression, and phylogeny. *Prog. Mol. Subcell. Biol.* **19:** 157–175.

Rodrigo, A. G., P. R. Bergquist, P. L. Bergquist, and R. A. Reeves. 1994. Are sponges animals? An investigation into the vagaries of phylogenetic interference. Pp. 47–54 in *Sponges in Time and Space,* R. W. M. v Soest, T. M. G. v. Kempen, and J. C. Braekman, eds. Balkema Press, Rotterdam.

Saitou, N., and M. Nei. 1987. The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4:** 406–425.

Scavo, L. M., A. R. Shuldiner, J. Serrano, R. Dashner, J. Roth, and F. de Pablo. 1991. Genes encoding receptors for insulin and insulin growth factor I are expressed in *Xenopus* oocytes and embryos. *Proc. Natl. Acad. Sci. USA* **88:** 6214–6218.

Schäcke, H., H. C. Schröder, V. Gamulin, B. Rinkevich, I. M. Müller, and W. E. G. Müller. 1994a. Molecular cloning of a receptor tyrosine kinase from the marine sponge *Geodia cydonium:* a new member of the receptor tyrosine kinase class II family in invertebrates. *Mol. Membr. Biol.* **11:** 101–107.

Schäcke, H., I. M. Müller, and W. E. G. Müller. 1994b. Tyrosine kinase from the marine sponge *Geodia cydonium:* the oldest member belonging to the receptor tyrosine kinase class II family. Pp. 201–211 in *Use of Aquatic Invertebrates as Tools for Monitoring of Environmental Hazards,* W. E. G. Müller, ed. Gustav Fischer Verlag, Stuttgart.

Seimiya, M., H. Ishiguro, K. Miura, Y. Watanabe, and Y. Kurosawa. 1994. Homeobox-containing genes in the most primitive Metazoa, the sponge. *Eur. J. Biochem.* **221:** 219–225.

Shier, P., and V. M. Watt. 1989. Primary structure of a putative receptor for a ligand of the insulin family. *J. Biol. Chem.* **264:** 14605–14608.

Suga, H., K. Kuma, N. Iwabe, N. Nikoh, K. Ono, M. Koyanagi, D. Hoshiyama, and T. Miyata. 1997. Intermittent divergence of the protein kinase family during animal evolution. *FEBS Lett.* **412:** 540–546.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22:** 4673–4680.

Ullrich, A., J. R. Bell, E. Y. Chen, R. Herrera, L. M. Petruzzelli, T. J. Dull, A. Gray, L. Coussens, Y. C. Liao, M. Tsubokawa, A. Mason, P. H. Seeburg, C. Grunfeld, O. M. Rosen, and J. Ramachandran. 1985. Human insulin receptor and its relationship to the tyrosine kinase family of oncogenes. *Nature* **313:** 756–761.

Ullrich, A., A. Gray, A. W. Tam, T. Yang-Feng, M. Tsubokawa, C. Collins, W. Henzel, T. Le Bon, S. Kathuria, E. Chen, S. Jacobs, U. Francke, J. Ramachandran, and Y. Fujita-Yamaguchi. 1986. Insulin-like growth factor I receptor primary structure: comparison with insulin receptor suggests structural determinants that define functional specificity. *EMBO J.* **5:** 2503–2512.

Valentine, J. W. 1994. The Cambrian explosion. Pp. 401–412 in *Early Life on Earth,* S. Bengtson, ed. Columbia University Press, New York.

Valentine, J. W., D. H. Erwin, and D. Jablonski. 1996. Developmental evolution of metazoan bodyplan: the fossil evidence. *Dev. Biol.* **173:** 373–381.

Zuckerkandl E., and L. Pauling. 1965. Evolutionary divergence and convergence in proteins. Pp. 97–166 in *Evolving Genes and Proteins.* V. Bryson and H. J. Vogel, eds. Academic Press, New York.